

Design Theory of Balanced Transistor Amplifiers

By K. KUROKAWA

(Manuscript received May 7, 1965)

This paper discusses the expected characteristics of balanced transistor amplifiers with symmetrical directional couplers.

Provided that pairs of transistors with similar characteristics can be selected from a given distribution, the input and output matches obtained with the balanced configuration are satisfactory over a ± 10 per cent bandwidth with simple one-section lumped-constant LC directional couplers and over a ± 40 per cent bandwidth (1.2 octaves) with one section distributed $\lambda/4$ couplers. For single-stage amplifiers, the decrease in gain is less than 0.1 db and the phase nonlinearities introduced by the couplers are about $\pm 0.15^\circ$ and $\pm 0.6^\circ$, respectively, over the same bandwidths.

The requirements on the terminations which are connected to the couplers to absorb the transistor reflections are not stringent: VSWR's less than 1.4 should be acceptable. The noise measure of balanced amplifiers is calculated to be a weighted average of the noise measures of the two component amplifiers, plus a small term which vanishes when the couplers have 3-db coupling and the component amplifiers have identical gains. Gain compression takes place at a 3-db higher signal level compared with conventional single-ended designs, and the expected improvement in the third-order intermodulation is 9 db on the average.

In the final section, the cascade connection of identical balanced amplifiers is discussed. With typical microwave transistors, the input and output return losses for a multistage amplifier should be about 4.5 db worse than those for the individual single-stage amplifiers of which it is composed. The gain ripple introduced by the interactions between stages is also investigated in detail.

I. INTRODUCTION

In a previous paper,¹ the principles and experimental results of an L-band balanced transistor amplifier have been discussed in which each

stage consists of two electrically similar transistors whose inputs and outputs are combined through 3-db directional couplers. Due to wide distributions in the characteristics of present microwave transistors, simultaneous realization of flat gain and good impedance matching is difficult to obtain with conventional single-ended designs unless, for instance, isolators are employed. On the other hand, as long as pairs of transistors with similar characteristics can be chosen from a given distribution, the balanced design offers good input and output impedance matches as well as smooth gain and phase characteristics, all simultaneously. Since the impedance matches are important in microwave systems, the balanced design will be useful for some time, until the distribution of transistor characteristics becomes so tight that a conventional single-ended design can easily provide good matches and smooth gain simultaneously.

While the theory given in the paper mentioned above should be adequate for general purposes, it may not be satisfactory for the actual design of balanced transistor amplifiers. The theory neglected interactions between the reflections which were introduced to explain the mismatches at the input and output ports of the transistors. It also assumed ideal 3-db coupling of the directional couplers for the entire frequency band of interest. The former shortcoming can be avoided by employing scattering matrices in the discussion. This paper is intended to supplement the previous one by presenting an improved theory which enables us to discuss the effect of the coupling variation on amplifier characteristics without resorting to too complicated mathematics. The noise performance and intermodulation characteristics are also included. In the final section, interactions between stages — when connected in cascade — are discussed in detail. Although wider bandwidths may be obtained by using couplers having characteristics slightly different from one another (e.g. stagger-tuning), for simplicity this paper assumes the use of identical couplers for both single and multistage amplifiers.

II. REVIEW OF DIRECTIONAL COUPLER PRINCIPLES

A directional coupler is a matched four-port network with zero coupling between conjugate ports. Let us consider a symmetrical directional coupler with two planes of symmetry as shown in Fig. 1. Because of the symmetry and by definition, the scattering matrix must have the form

$$S = \begin{bmatrix} 0 & \alpha & \beta & 0 \\ \alpha & 0 & 0 & \beta \\ \beta & 0 & 0 & \alpha \\ 0 & \beta & \alpha & 0 \end{bmatrix} \quad (1)$$

If a network is lossless, the scattering matrix has to satisfy

$$S^+ S = I \quad (2)$$

where $+$ indicates the transposed conjugate matrix and I the unit matrix. From this, two constraints between α and β of a lossless symmetrical coupler are obtained:

$$|\alpha|^2 + |\beta|^2 = 1, \quad \alpha^* \beta + \beta^* \alpha = 0. \quad (3)$$

The first equation specifies a relation between magnitudes and the second one indicates that α and β must be 90° out of phase. Thus, α and β must be expressible in terms of two real quantities t and φ ,

$$\alpha = \sqrt{1 - t^2} j \exp(-j\varphi), \quad \beta = t \exp(-j\varphi). \quad (4)$$

That is, if a unit wave is incident on port 1, the output waves from ports 2 and 3 are given by $\sqrt{1 - t^2} j \exp(-j\varphi)$ and $t \exp(-j\varphi)$, respectively, and port 4 has no output.

There is a class of symmetrical junctions which acts as directional couplers independent of the frequency. Let us consider two of them as examples: a lumped constant directional coupler, and a distributed

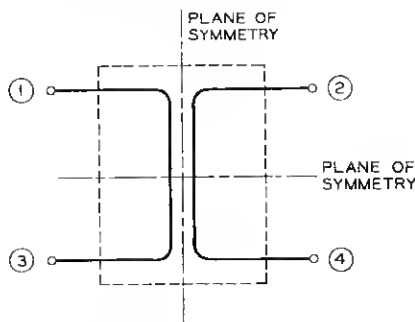


Fig. 1 — Schematic diagram of symmetrical directional coupler.

transmission line directional coupler $\lambda/4$ long at center frequency of operation.

For the lumped-constant directional couplers (Appendix) with a common inductance L from ports 1 and 2 to ports 3 and 4 and with a capacitance C between ports 1 and 2 or 3 and 4,

$$t = \frac{1}{\sqrt{1 + \zeta^2}}, \quad \varphi = \tan^{-1} \zeta \quad (5)$$

where $\zeta = \omega L/Z_o$ and the characteristic impedance $Z_o = \sqrt{L/C}$. When $\zeta = 1$, $t^2 = 1 - t^2 = 0.5$ and 3-db coupling is obtained. Figs. 2 and 3 show t and φ vs the normalized frequency f/f_o where f_o is the frequency for ζ being 1.

For the distributed directional coupler²

$$t = \sqrt{\frac{1 - k^2}{1 - k^2 \cos^2 \theta}}, \quad \varphi = \tan^{-1} \frac{1}{\sqrt{1 - k^2}} \tan \theta \quad (6)$$

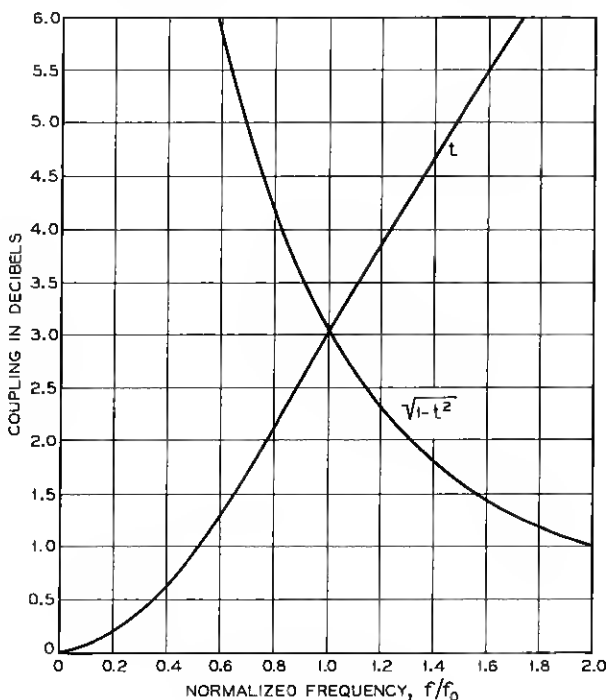


Fig. 2—Coupling vs normalized frequency of lumped-constant LC directional coupler.

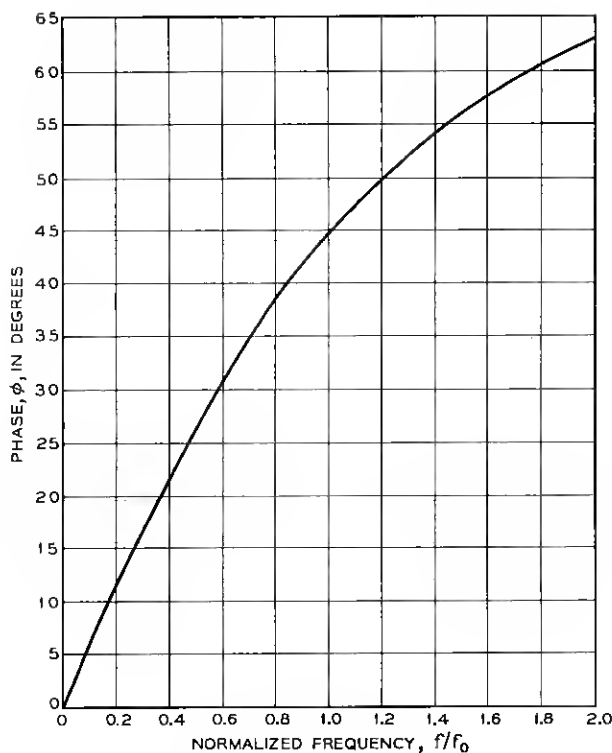


Fig. 3—Phase vs normalized frequency of lumped-constant LC directional coupler.

where k is the coupling factor in the theory of coupled transmission lines and θ is the electrical length of the coupled region. In terms of the even- and odd-mode characteristic impedances Z_{oe} and Z_{oo} , the characteristic impedance Z_o and the coupling factor k are given by

$$Z_o = \sqrt{Z_{oe}Z_{oo}}, \quad k = \frac{Z_{oe} - Z_{oo}}{Z_{oe} + Z_{oo}}, \quad (7)$$

respectively. Figs. 4 and 5 give t and ϕ vs the normalized frequency f/f_o and f_o is the frequency for which $\theta = \pi/2$ or 90° .

III. SCATTERING MATRIX OF ONE-STAGE BALANCED AMPLIFIER

Let us consider the configuration shown in Fig. 6 where two transistors, a and b , are connected by two directional couplers in which ports 3 and 4 are crossed over (as compared to Fig. 1). Due to the lack

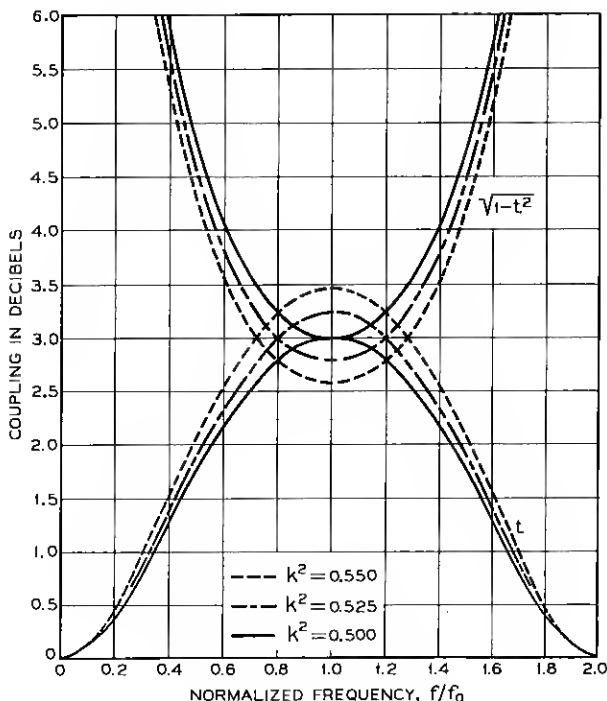


Fig. 4 — Coupling vs normalized frequency of one-section distributed couplers. k : coupling factor.

of coupling between conjugate arms (1-4 and 2-3 in Fig. 1) of the couplers, the components of the over-all scattering matrix between ports 1 and 2, are easily calculated. They are:

$$\begin{aligned}
 S_{11} &= e^{-2j\varphi} [t^2 S_{11}(a) - (1 - t^2) S_{11}(b)] \\
 S_{21} &= j e^{-j2\varphi} t \sqrt{1 - t^2} [S_{21}(a) + S_{21}(b)] \\
 S_{12} &= j e^{-j2\varphi} t \sqrt{1 - t^2} [S_{12}(a) + S_{12}(b)] \\
 S_{22} &= e^{-2j\varphi} [t^2 S_{22}(b) - (1 - t^2) S_{22}(a)].
 \end{aligned} \tag{8}$$

The subscripts 1 and 2 refer to the input and output ports, respectively, and $S_{11}(a)$, $S_{11}(b)$ etc., are the scattering matrix components of transistors a and b including their surrounding circuits, i.e., the scattering matrix components of the component amplifiers a and b , respectively. When the coupling is 3 db ($t^2 = 0.5$) and the two component amplifiers are similar in their characteristics,

$$S_{ij}(a) \approx S_{ij}(b)$$

and hence

$$|S_{11}| \approx 0, \quad |S_{22}| \approx 0, \quad S_{21} \approx je^{-j\theta} S_{21}(a) \approx je^{-j\theta} S_{21}(b).$$

This means that the input and output ports of the balanced amplifier are well matched and the gain is approximately equal to that of either

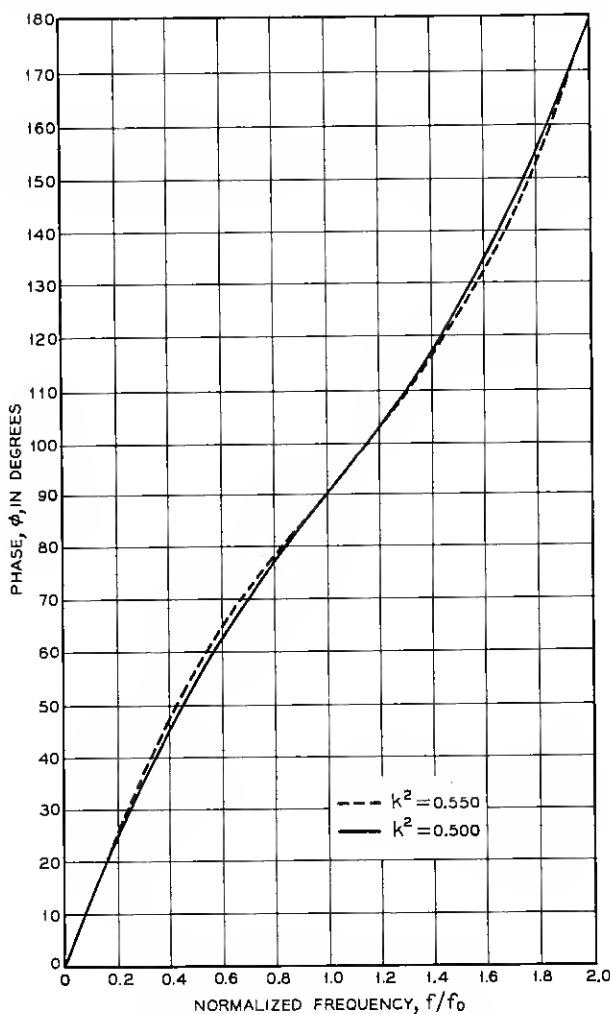


Fig. 5 — Phase vs normalized frequency of one-section distributed couplers.

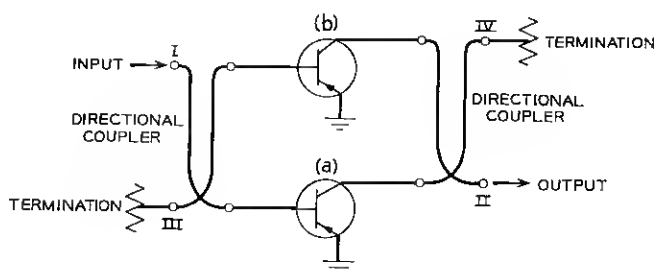


Fig. 6 — Schematic diagram of one-stage balanced transistor amplifier.

component amplifier. Because of the term, $\exp(-j2\phi)$, the phase of the balanced amplifier is affected by the 3-db couplers. This will be discussed in detail in Section V. The reflections from the transistors are absorbed in the terminations connected to the couplers as we shall discuss in Section VI.

When the transistors are dissimilar, but with the coupling still 3 db,

$$|S_{11}| = \frac{1}{2} |S_{11}(a) - S_{11}(b)|, \quad |S_{22}| = \frac{1}{2} |S_{22}(a) - S_{22}(b)|$$

and

$$|S_{21}| = \frac{1}{2} |S_{21}(a) + S_{21}(b)|.$$

That is, the input and output reflections are reduced to half of the vector differences of the corresponding reflections of the transistors, and the gain is given by the vector average of the two gains.

In this section, the coupling of the directional couplers has been assumed to be 3 db. In Section IV, we shall investigate the effect on the amplifier characteristics when the coupling is not 3 db.

IV. COUPLING OF THE DIRECTIONAL COUPLERS AND AMPLIFIER CHARACTERISTICS

First, let us assume that the two component amplifiers are similar in their characteristics. Then, from (8), we have

$$\begin{aligned} |S_{11}| &\approx |2t^2 - 1| |S_{11}(a)| \\ |S_{22}| &\approx |2t^2 - 1| |S_{22}(a)| \\ |S_{21}| &\approx |2t\sqrt{1-t^2}| |S_{21}(a)|. \end{aligned} \tag{9}$$

If $|2t^2 - 1|$ is given in terms of loss in db and $|S_{11}(a)|$ or $|S_{22}(a)|$

in terms of return loss in db, then the addition of these figures gives the corresponding return loss of the balanced amplifier in db. Similarly, if the gain of the component amplifiers, $|S_{21}(a)|$, is expressed in db and $|2t\sqrt{1-t^2}|$ in terms of loss in db, then the difference of these two figures gives the gain of the balanced amplifier in db. Fig. 7 shows the losses $|2t^2 - 1|$ and $|2t\sqrt{1-t^2}|$ in db vs the coupling loss t in db. From Fig. 7 we see, for instance, that if the input and output VSWR's of the component amplifiers are better than 2 (return loss 10 db), the coupling loss t can deviate as much as -0.4 db and $+0.5$ db from 3 db before the VSWR's of the balanced amplifier become worse than 1.07 (return loss 30 db) and that the decrease of the gain due to the directional couplers from that of the component amplifier is less than 0.1 db. The above deviation allows ± 10 per cent and ± 40 per cent frequency bandwidths for the lumped-constant and the distributed ($k^2 = 0.550$) couplers, respectively.

Next, let us consider the case where the two component amplifiers have different characteristics. In this case, from (8), we have

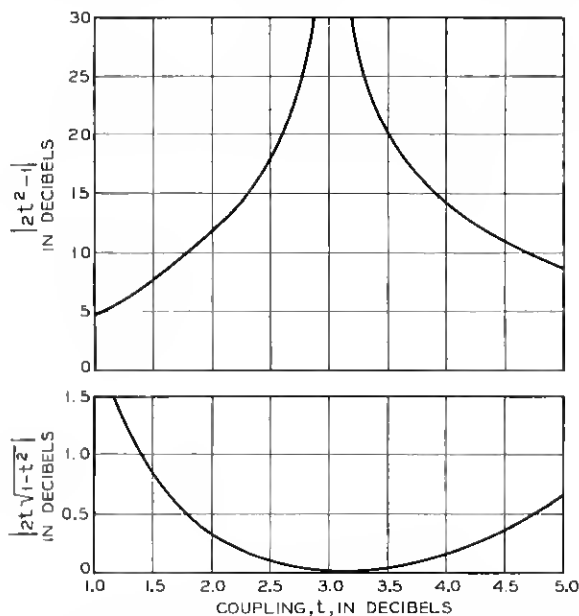


Fig. 7 — Improvement in return loss $|2t^2 - 1|$ in db and decrease in gain $|2t\sqrt{1-t^2}|$ in db due to the directional couplers in balanced design.

$$\begin{aligned}
|S_{11}| &= \left| (2t^2 - 1) \frac{S_{11}(a) + S_{11}(b)}{2} + \frac{S_{11}(a) - S_{11}(b)}{2} \right| \\
|S_{22}| &= \left| (2t^2 - 1) \frac{S_{22}(a) + S_{22}(b)}{2} - \frac{S_{22}(a) - S_{22}(b)}{2} \right| \quad (10) \\
|S_{21}| &= |2t \sqrt{1 - t^2}| \left| \frac{S_{21}(a) + S_{21}(b)}{2} \right|.
\end{aligned}$$

The gain expression in (10) is the same as that in (9) except $S_{21}(a)$ is replaced by the mean vector between $S_{21}(a)$ and $S_{21}(b)$. Therefore, using the magnitude of the mean vector and Fig. 7, the expected gain of the balanced amplifier can be easily calculated. The reflections $|S_{11}|$ and $|S_{22}|$ have two terms each: one becomes small when the coupling approaches 3 db and the other is independent of the coupling. The magnitude of the first term can be evaluated by using Fig. 7 as we have done before. Now, however, the mean reflection from the two amplifiers is used instead of the same reflection from the component amplifiers. The magnitude of the second term is half of the difference between the reflections from the two component amplifiers. In order to get the resultant reflection, however, a vectorial addition of these two terms is necessary. For more clear understanding of the situation, the following is an additional way of viewing the same problem.

Rearranging the first two equations in (10), we have

$$\begin{aligned}
|S_{11}| &= |(2t^2 - 1)S_{11}(a) - (1 - t^2)\Delta_1| \\
|S_{22}| &= |(2t^2 - 1)S_{22}(a) + t^2\Delta_2|,
\end{aligned} \quad (11)$$

where Δ_1 and Δ_2 are given by

$$\Delta_1 = S_{11}(b) - S_{11}(a), \quad \Delta_2 = S_{22}(b) - S_{22}(a).$$

Suppose that the coupling t and $S_{11}(a)$ are specified and that the resultant reflection $|S_{11}|$ is required to be smaller than a certain magnitude, $|S_{11}|_{\max}$. Then, referring to Fig. 8, the tip of $-(1 - t^2)\Delta_1$, drawn from $(2t^2 - 1)S_{11}(a)$, must lie inside a circle centered at the origin and of radius $|S_{11}|_{\max}$. Expanding the figure by a factor of $-1/(1 - t^2)$, the tip of Δ_1 , drawn from $-(2t^2 - 1)S_{11}(a)/(1 - t^2)$, is seen to be inside a circle centered at the origin and of radius $|S_{11}|_{\max}/(1 - t^2)$. Now translate the whole figure until the tail of Δ_1 falls on the point $S_{11}(a)$. Then, $S_{11}(b)$ is seen to be inside a circle centered at $t^2S_{11}(a)/(1 - t^2)$ and of radius $|S_{11}|_{\max}/(1 - t^2)$ on the Smith chart. Similarly, if the maximum allowable value of $|S_{22}|$ is given by $|S_{22}|_{\max}$, $S_{22}(b)$ must be inside a circle centered at $(1 - t^2)S_{22}(a)/t^2$ and of radius $|S_{22}|_{\max}/t^2$.

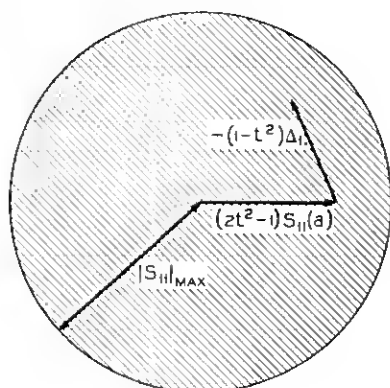


Fig. 8 — Vector diagram for the relation $|S_{11}| \leq |S_{11}|_{\max}$.

As t^2 increases, the area in which $S_{11}(b)$ should be located also increases; however, the area for $S_{22}(b)$ decreases. The best compromise is obtained when the coupling is 3 db. Here, one might ask the following question: If there is no requirement for $|S_{22}|$, should one make t^2 as large as possible in order to achieve the required matching more easily? The answer is, generally, no. Since $S_{11}(a)$ is usually larger than $|S_{11}|_{\max}$, as t^2 approaches 1, the center $t^2 S_{11}(a)/(1 - t^2)$ moves faster than the increase in the radius $|S_{11}|_{\max}/(1 - t^2)$. The circle therefore ceases to cover the area where the transistor distribution for S_{11} is dense and it becomes harder to find a proper transistor which gives the required $S_{11}(b)$. In this argument, if $S_{11}(a)$ is always smaller than $|S_{11}|_{\max}$, t^2 could obviously be 1. However, if the reflection from the component amplifier a is already smaller than the required value there is no reason for using the balanced configuration and a second amplifier. A similar argument holds for the output match as well.

V. PHASE LINEARITY

It is obvious from (8) that the phase linearity of the balanced amplifier depends on the phase characteristics of the directional couplers as well as on the phase linearity of the component amplifiers. The φ 's of the couplers were discussed in Section II and are shown in Figs. 3 and 5. From these figures, the phase nonlinearity introduced by the directional couplers can be estimated. However, in precise applications the over-all phase linearity required is often within a few degrees and if several stages in cascade are employed to obtain a desired gain, the phase linearity required for each stage would be within a fraction of one degree. In such a case, the Taylor expansion of 2φ around $f = f_0$ should give a

better estimate of the nonlinearity introduced by the couplers. For the lumped constant directional couplers,

$$\begin{aligned} 2\varphi &= \frac{1}{2}\pi + \chi - \frac{1}{2}\chi^2 + \frac{1}{6}\chi^3 + \cdots & (2\varphi, \text{ in radians}) \\ &= 90 + 57.3\chi - 28.7\chi^2 + 9.6\chi^3 + \cdots & (2\varphi, \text{ in degrees}) \end{aligned} \quad (12)$$

where

$$\chi = (f - f_o)/f_o. \quad (13)$$

For the distributed couplers $\lambda/4$ long at f_o ,

$$\begin{aligned} 2\varphi &= \pi + \pi\sqrt{1 - k^2}\chi + \frac{k^2\sqrt{1 - k^2}}{12}\pi^3\chi^3 \\ &\quad + \frac{k^2(3k^2 - 1)\sqrt{1 - k^2}}{240}\pi^5\chi^5 + \cdots \quad (2\varphi, \text{ in radians}) \\ &= 180 + 180\sqrt{1 - k^2}\chi + 148k^2\sqrt{1 - k^2}\chi^3 \\ &\quad + 73k^2(3k^2 - 1)\sqrt{1 - k^2}\chi^5 + \cdots \quad (2\varphi, \text{ in degrees}). \end{aligned} \quad (14)$$

For instance, the deviations of the lumped-constant couplers from phase linearity at $\chi = f/f_o - 1 = 0.1, 0.2$ and 0.3 are $0.3^\circ, 1.2^\circ$ and 2.6° respectively, and of the order of $0.04^\circ, 0.3^\circ$ and 1° for the distributed couplers. These deviations are measured from the straight line tangential to the 2φ curve at $f = f_o$. If the reference line is redrawn so that the maximum deviation becomes the smallest, these figures will be about one-half of those mentioned above for the lumped-constant couplers and about one-fourth for the distributed couplers. Thus, over the bandwidths for which the input and output VSWR's are less than 1.07 as discussed before, the phase nonlinearities introduced are of the order of $\pm 0.15^\circ$ and $\pm 0.6^\circ$, respectively.

VI. EFFECT OF IMPERFECT TERMINATIONS

In order to investigate in detail the role of the terminations connected to the couplers, let us first consider the balanced amplifier as a four-port network rather than a two-port network as we have done so far. The ports are numbered by Roman numerals as shown in Fig. 6. Since there is no coupling between conjugate ports of the couplers, the scattering matrix of the four-port network can again be easily calculated. S_{11}, S_{21}, S_{12} , and S_{22} are the same as given in (8). S_{43} and S_{34} are equal to S_{21} and S_{12} respectively. S_{33} and S_{44} are the same as S_{11} and S_{22} , respectively, except that the component amplifier designations a and b are interchanged. The others are given by

$$\begin{aligned}
S_{13} &= S_{31} = je^{-2j\varphi}t\sqrt{1-t^2}[S_{11}(a) + S_{11}(b)] \\
S_{24} &= S_{42} = je^{-2j\varphi}t\sqrt{1-t^2}[S_{22}(a) + S_{22}(b)] \\
S_{23} &= e^{-2j\varphi}[t^2S_{21}(b) - (1-t^2)S_{21}(a)] \\
S_{32} &= e^{-2j\varphi}[t^2S_{12}(b) - (1-t^2)S_{12}(a)] \\
S_{14} &= e^{-2j\varphi}[t^2S_{12}(a) - (1-t^2)S_{12}(b)] \\
S_{41} &= e^{-2j\varphi}[t^2S_{21}(a) - (1-t^2)S_{21}(b)].
\end{aligned} \tag{15}$$

When a wave of unity power is incident to port I, it is split by the input coupler into two, t^2 and $(1-t^2)$, arriving at the component amplifiers a and b , respectively. The reflections there are given by $t^2 |S_{11}(a)|^2$ and $(1-t^2) |S_{11}(b)|^2$. Of these, only $|S_{11}|^2$ comes back to port I and the rest of the power

$$\begin{aligned}
t^2 |S_{11}(a)|^2 + (1-t^2) |S_{11}(b)|^2 - |S_{11}|^2 \\
= t^2(1-t^2) |S_{11}(a) + S_{11}(b)|^2
\end{aligned} \tag{16}$$

(which is exactly equal to $|S_{31}|^2$), goes to port III. Thus, most of the reflected power from the component amplifiers goes to port III and is absorbed there. A similar argument holds for the output port. These are the reasons why the balanced configuration gives good matches at both ends.

Next, let us investigate how critical the matches are for these terminations. Indicating the reflection coefficients of the terminations by r_3 and r_4 (the subscripts refer to port numbers) and drawing the signal flow graph as shown in Fig. 9, the reflection to port I, S_{11}' can be written down by inspection.

$$S_{11}' = S_{11} + \frac{r_3 S_{13} S_{31} (1 - r_4 S_{44}) + r_4 S_{14} S_{41} (1 - r_3 S_{33}) + r_3 r_4 (S_{13} S_{34} S_{41} + S_{14} S_{43} S_{31})}{1 - r_3 S_{33} - r_4 S_{44} - r_3 r_4 S_{43} S_{34} + r_3 r_4 S_{33} S_{44}}. \tag{17}$$

Neglecting higher order terms, S_{11}' can be approximated by

$$S_{11}' \approx S_{11} + r_3 S_{13} S_{31} + r_4 S_{14} S_{41}. \tag{18}$$

The first term on the right-hand side represents the reflection when $r_3 = r_4 = 0$, the second term the reflection due to r_3 and the last term due to r_4 . The neglected terms represent the contribution due to multiple reflections between the terminations and they are in general so small compared with the terms given above that their omission is readily justified.

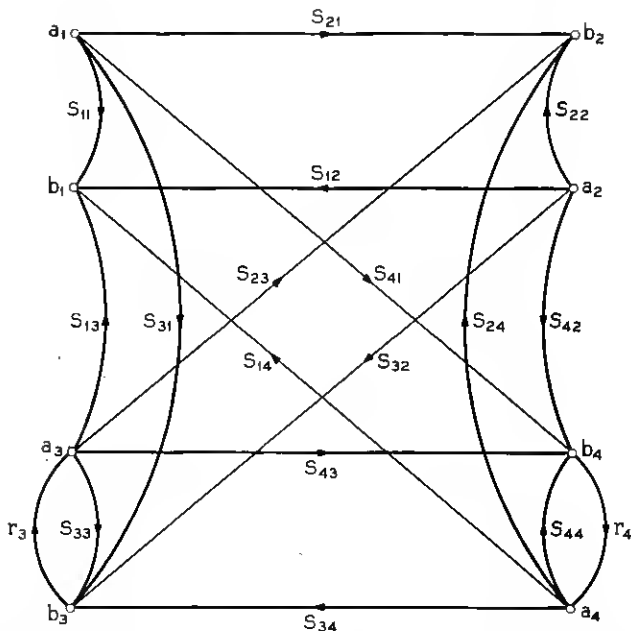


Fig. 9 — Signal flow graph for one-stage balanced amplifier with imperfect terminations.

Similarly, S_{22}' and S_{21}' are approximately given by

$$\begin{aligned} S_{22}' &\approx S_{22} + r_4 S_{24} S_{42} + r_3 S_{23} S_{32} \\ S_{21}' &\approx S_{21} + r_3 S_{23} S_{31} + r_4 S_{24} S_{41}. \end{aligned} \quad (19)$$

Typical values for the magnitude of the coefficients of r_3 and r_4 appearing in S_{11}' and S_{22}' are of the order of -20 db and those appearing in S_{21}' are of the order of -25 db or less. Therefore, if $|r_3|$ and $|r_4|$ are both less than -15 db ($\text{SWR} \leq 1.4$), then the effect of imperfect terminations on the amplifier characteristics must be negligibly small. In conclusion, the required matches for the terminations are in general not so stringent; SWR's of less than 1.4 are usually acceptable.

VII. NOISE PERFORMANCE

Noise performance of an amplifier is evaluated by the actual noise measure.³ It is defined by

$$M = \frac{F - 1}{1 - (1/G)} \quad (20)$$

where G is the transducer gain and F is the noise figure of the amplifier, including the contribution of the noise power originating in and reflected back to the output load. When the input and output of the amplifier are matched, a number of amplifiers with identical characteristics can be connected in cascade, making the total gain very high. The excess noise figure of the high gain amplifier is then given by M itself. However, when the input and output are not matched, we have only to insert isolators between the stages in order to reach the same interpretation for M . For each amplifier there is an optimum value, M_{opt} , of M which can be achieved by a lossless imbedding but cannot be surpassed by any passive transformation of the amplifier. The noise measure itself is a dimensionless number.

Now, let us consider the balanced amplifier. The terminations III and IV connected to ports III and IV, respectively, are assumed to be matched. Furthermore, the noise temperatures of the terminations as well as of the load are assumed to be 290°K . For the ideal case where the coupling of the directional couplers is 3 db and the characteristics of the two component amplifiers are identical, the noise originating in each component amplifier is split into two. Only a half of the total power goes to the output load, with the other half going to termination IV. Since there are two component amplifiers, the output load receives the same noise power as in the single-ended design. The noise originating in termination III is amplified but absorbed in termination IV and none of it comes out to the load. The noise originating in the load and reflected back from the component amplifiers goes to termination IV and does not come back to the load. However, noise power originating in termination IV goes into the load. This power is exactly equal to the noise power originating in and reflected back to the load ($T = 290^\circ$) in the single-ended design. As a result, the noise measure M of the one-stage balanced amplifier in this ideal case is equal to the noise measure of either component amplifier.

When a transistor is unconditionally stable, by inserting a proper lossless circuit at the input and a matching circuit at the output, the optimum value M_{opt} for the transistor can be achieved. Therefore, the component amplifiers can have M_{opt} in this way, which means that the balanced amplifier can also give M_{opt} . It is worth noting that this realization of M_{opt} does not deteriorate the input matching of the balanced amplifier. In general, this is not the case for single-ended designs.

Next, let us consider the case where the coupling is not necessarily 3 db and the component amplifiers have different characteristics. The assumptions for the terminations and the load remain the same as be-

fore. The excess noise output to the load is given by

$$\begin{aligned} N = & \{ (F_a - 1)kTB |S_{21}(a)|^2 - kTB |S_{22}(a)|^2 \} (1 - t^2) \\ & + \{ (F_b - 1)kTB |S_{21}(b)|^2 - kTB |S_{22}(b)|^2 \} t^2 \quad (21) \\ & + kTB |S_{23}|^2 + kTB |S_{24}|^2 + kTB |S_{22}|^2, \end{aligned}$$

where F_a and F_b are the noise figures of the component amplifiers a and b , respectively. The first term on the right-hand side of (21) represents the output noise originating in component amplifier a , the second one in component amplifier b , the third one in termination III, the fourth one in termination IV, and the last one represents the noise originating in and reflected back to the load. Combination of (8), (15) and (21) together with (20) gives the noise measure M of the balanced amplifier as follows:

$$M = \frac{M_a(|S_{21}(a)|^2 - 1)(1 - t^2) + M_b(|S_{21}(b)|^2 - 1)t^2 + |S_{23}|^2}{|S_{21}|^2 - 1}. \quad (22)$$

Thus, M is a weighted average of the noise measures M_a and M_b of the component amplifiers plus a small term which comes in because of termination III. To make this additional term small, from (15) $t^2 S_{21}(b)$ and $(1 - t^2) S_{21}(a)$ should be close to each other. When the coupling is 3 db, this means that the two component amplifiers should have approximately equal gains. Thus, we see that for the balanced design, a pair of transistors should be selected on the basis of close S_{11} , S_{22} and S_{21} ; the first two being necessary for good matches and the last for low noise (although, in practice, the last requirement is not stringent at all).

VIII. GAIN COMPRESSION AND INTERMODULATION

Since each transistor handles only one-half of the signal power, it begins to saturate at a 3-db higher signal level thus improving the gain compression and intermodulation characteristics. The type of intermodulation of most concern in broadband amplifiers with multiple frequency channels is usually one in which two strong signals of frequency f_A and f_B produce third order intermodulation signals at frequencies $2f_A - f_B$ and $2f_B - f_A$, also within the passband of the amplifier, where they might interfere with wanted weak signals. Since the signal level to each transistor is 3-db lower, the third order intermodulation output from each transistor must be 9-db lower. Thus, if the two intermodulation outputs are in phase, a resultant output of 6-db below that of the single-ended design is expected for the balanced amplifier.

However, the magnitude of the third order intermodulation output varies between transistors at microwave frequencies, even if the transistor characteristics for the signal frequency are quite similar. This suggests that the phase of the intermodulation output might also be random. In this case, the resultant output of the balanced amplifier is expected to be 9-db lower on the average, instead of 6 db. This conclusion is strongly supported by experimental results on 18 different pairs of transistors.⁴

IX. CASCADE CONNECTION

So far we have discussed only single-stage balanced amplifiers. In this section, let us consider the interactions between stages when connected in cascade. Since the outgoing wave of the n th stage output port is equal to the incoming wave of the $(n + 1)$ th stage input port, the signal flow graph of a multistage amplifier becomes something like Fig. 10, where $S_{11}(n)$ etc., are the scattering matrix components of the

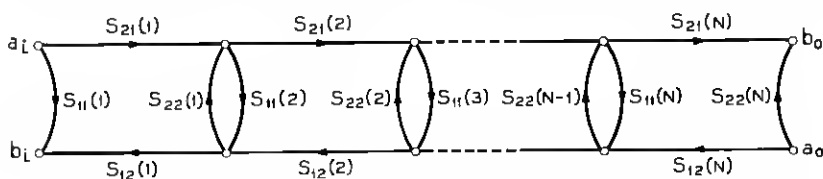


Fig. 10 — Signal flow graph of multistage amplifier.

n th stage and a_i , b_i , a_o and b_o are the incident and outgoing waves at the input (subscript i) and the output (subscript o) ports of the cascaded amplifier, respectively. Inspecting this signal flow graph, the components of the over-all scattering matrix can be obtained. First, taking a three-stage amplifier as an example, let us consider the input reflection S_{11} of the amplifier. It is given by

$$S_{11} = S_{11}(1) + \frac{S_{11}(2)S_{21}(1)S_{12}(1)}{\Delta} \{1 - S_{22}(2)S_{11}(3)\} + \frac{S_{11}(3)S_{21}(1)S_{12}(1)S_{21}(2)S_{12}(2)}{\Delta}, \quad (23)$$

where

$$\begin{aligned} \Delta = & 1 - S_{22}(1)S_{11}(2) - S_{22}(2)S_{11}(3) \\ & - S_{22}(1)S_{21}(2)S_{11}(3)S_{12}(2) \\ & + S_{22}(1)S_{11}(2)S_{22}(2)S_{11}(3). \end{aligned} \quad (24)$$

Since $|S_{22}(n)S_{11}(n+1)|$ is small compared with unity for most practical balanced amplifiers and since $\rho = |S_{21}(n)S_{12}(n)|$ is of the order of 0.4 for typical transistors, S_{11} can be approximated by

$$S_{11} \approx S_{11}(1) + S_{11}(2)S_{21}(1)S_{12}(1) + S_{11}(3)S_{21}(1)S_{12}(1)S_{21}(2)S_{12}(2). \quad (25)$$

The first, second, and third terms on the right-hand side of (25) represent the contributions to S_{11} by the reflections from the first, second and third stages respectively. The effect of later stages is seen to be reduced by the buffer action of the previous stages indicated by $S_{21}(n)S_{12}(n)$. For instance, when $\rho = |S_{21}(n)S_{12}(n)|$ is approximately equal to 0.4, the contribution of the third stage to the over-all mismatch is reduced to only $\rho^2 = 0.16$ times the original reflection. When the frequency is changed, the phase of $S_{21}(n)S_{12}(n)$ as well as that of $S_{11}(n)$ changes. Therefore, the vectors representing the successive terms on the right-hand side of (25) are expected to rotate with successively increasing speeds. At some frequencies, they tend to cancel each other and at other frequencies they tend to add up. Thus, if the reflection of each stage is of the same order of magnitude and $\rho = |S_{21}(n)S_{12}(n)| \approx 0.4$, then for the three-stage amplifier, $1 + \rho + \rho^2 \approx 1.56$ times as large reflection as the single stage should be expected (or 4 db worse return loss). Similarly, for a multistage amplifier with a large number of stages, $1 + \rho + \rho^2 + \dots = 1/1 - \rho \approx 1.67$ times as large reflection should be anticipated (or 4.5 db worse return loss). The output reflection S_{22} can be discussed in a similar manner.

Next, let us consider the gain S_{21} . S_{21} of the three-stage amplifier is given by

$$S_{21} = \frac{S_{21}(1)S_{21}(2)S_{21}(3)}{\Delta}, \quad (26)$$

where Δ is given by (24).

Since factors other than the effect of Δ being different than one were dominant, Δ could be approximated by unity for the discussion of S_{11} . However, for discussing S_{21} , Δ has to be investigated in detail. When each stage is well matched, Δ is unity and the over-all gain is the product of the gain of each stage as expected. The effect on the gain of the interaction between stages comes from the various terms in Δ . The second and third terms on the right-hand side of (24) represent the effect of the interaction between the adjacent stages. The fourth term shows the effect of the interaction between the first and the third stages through

some huffer action of the second stage. The last term gives the higher order interaction which for well designed balanced amplifiers can usually be neglected. Since the phases as well as the magnitudes of the $S_{ij}(n)$'s change with frequency, the interactions between stages introduce ripples in the gain vs frequency curve. When $|S_{11}(n)|$ and $|S_{22}(n)|$ are of the order of 0.1 (or 20-db return loss) the magnitudes of the ripples introduced by the second and third terms in Δ are of the order of ± 0.1 db. Therefore, the expected value of the resultant is ± 0.14 db (or ± 0.2 db for the worst case). The magnitude of the ripple due to the fourth term is smaller by the factor of $|S_{21}(2)S_{12}(2)|$. However, the phase of $S_{21}(2)S_{12}(2)$ increases the speed with which the vector rotates with frequency. This rapid variation is sometimes troublesome. To reduce the repetition rate of the gain variation with frequency, the equivalent electrical length of each stage should be made as small as possible. Also, transistors with high reverse loss help reduce the magnitude of the rapid ripple.

For an N -stage amplifier, the corresponding Δ includes $N - 1$ terms representing the interactions between the adjacent stages, $N - 2$ terms representing those between the n th and $n + 2$ nd stages and so forth — with additional terms representing various higher order interactions. The contribution from each group to the gain ripple is proportional to $\sqrt{N - 1}$, $\sqrt{N - 2}$ ρ and so forth, provided that all stages have similar characteristics. Although the speed of the rotation with frequency increases successively, the magnitude of the vector representing each group diminishes rapidly and practically no interaction between the stages beyond 3 stages away from each other can be observed when $\rho \approx 0.4$.

Since each balanced stage is, in practice, well matched at both ends, the noise performance and intermodulation characteristics of a multi-stage amplifier with identical stages are clear from the discussions of the single-stage amplifier. However, because the main noise contribution comes from the first few stages and the contribution to the compression or intermodulation comes from the last few stages, it may be advisable not to use an identical design for all stages. Instead, the first few stages can be designed for best noise performance and the last few stages for best compression characteristics. This can usually be done by changing only the transistor de bias circuit.

In large scale production, when identical circuits are to be used for the first several stages of each amplifier, the following procedure of selecting transistor pairs gives the best noise performance on the average. First, obtain the actual noise measure M of all transistors in a standard component amplifier and classify them into several groups of increasing

M , each group containing twice as many transistors as the number of amplifiers to be built. Then select pairs of transistors from each group separately on the basis of similar scattering matrices and use first, each pair from the best group (lowest M) in the first stage of each amplifier, next use those from the second group in the second stage and so forth. The pairs from the last group with poor M 's must be used in the later stages, whose noise contributions are insignificant. A similar procedure can be applied to the selection of transistor pairs for best compression characteristics. Here, of course, each pair from the best compression group is used in the last stage of each amplifier.

X. ACKNOWLEDGMENTS

Acknowledgments are due to R. S. Engelbrecht who originally suggested the balanced transistor amplifiers and supervised this work, and to K. M. Eisele and L. D. Gardner without whose cooperation the whole project of L-band balanced amplifiers would not have been successful.

APPENDIX

Theory of Symmetrical Directional Couplers

Since there is no literature readily available on the lumped constant directional coupler discussed in the text, this appendix is prepared to explain its principle from a slightly broader point of view. The theory to be presented is originally due to H. Seidel. It was developed during his association with Merrimac Research and Development, Inc., and is used in the design of their low frequency directional couplers.

For convenience, let us call two two-port networks "oppositely reflective" with respect to each other when they have identical scattering matrices except for opposite signs of their diagonal components. Now, let us consider the symmetrical network shown in Fig. 11, and apply incident waves of an even mode to ports 1 and 2, i.e., waves with the same amplitude and phase. Because of this symmetry, the actual (four-port) network can be considered as a two-port network acting on the incident mode. We thus have some reflection r_e of the even mode from ports 1 and 2, and a transmission t_e of the even mode to ports 3 and 4. Next, suppose that we apply incident waves of an odd mode to ports 1 and 2, i.e., waves with the same amplitude, but 180° out of phase. Again, because of the symmetry, the actual network acts as a two-port network to the incident mode and we have some reflection r_o and trans-

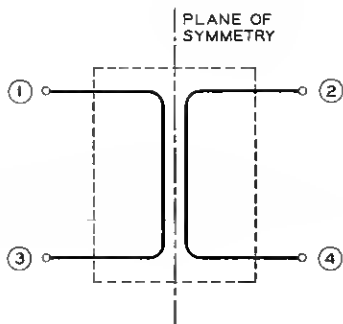


Fig. 11 — Symmetrical directional coupler with only one plane of symmetry.

mission t_o of the odd mode from ports 1 and 2 and from ports 3 and 4, respectively. With this much preparation, let us present the following theorem.

Theorem I: If the two two-port networks, presented by a symmetrical four-port network to its even and odd modes, are oppositely reflective, then the symmetrical four-port network is a directional coupler. (The converse is not necessarily true.)

The proof of this theorem is as follows. By definition, $r_e = -r_o$ and $t_e = t_o$. Therefore, let us define α and β by

$$\alpha = r_e = -r_o, \quad \beta = t_e = t_o.$$

Suppose that a unit wave is incident at port 1. This can be considered as a superposition of even and odd modes incident at ports 1 and 2 with amplitudes of one half each. This fact can be expressed in matrix form

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

where the upper and lower rows represent the waves on the left- and right-hand sides of the vertical plane of symmetry in Fig. 11, respectively. The reflection from ports 1 and 2 is therefore given by

$$\frac{r_e}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{r_o}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha \end{bmatrix}.$$

This means that port 1 has no reflection and port 2 has an output α . Similarly, the transmission to ports 3 and 4 is given by

$$\frac{t_e}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{t_o}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \beta \\ 0 \end{bmatrix}.$$

This tells us that port 3 has an output β but no output appears at port 4. In other words, when a unit wave is incident to port 1, port 1 is matched, ports 2 and 3 have output waves α and β , respectively, and port 4 has no output. A similar argument holds for a unit wave incident at any other port of the symmetrical four-port network. This means that each port is matched and there is no coupling between conjugate ports. Thus, if a symmetrical four-port network is oppositely reflective to its even and odd modes in the sense discussed above, then it is a directional coupler and the theorem is proved.

The next theorem is useful for searching possible structures of directional couplers.

Theorem II: If two two-port networks with real generator and load immittances are dual in their normalized form with respect to the generator immittances, then the two-port networks are oppositely reflective independent of the frequency.

To make the meaning of the theorem clear and the proof easy, let us consider a simple example as shown in Fig. 12. For later use, the normalized load immittances are assumed to be unity in Fig. 12; however, this assumption is not necessary for the present discussion. The duality is satisfied when the normalized inductance l is equal to the normalized capacitance c . The theorem asserts that the networks inside the dotted lines are oppositely reflective at all frequencies. However, this is obvious from the following consideration. The normalized input (or output) impedance of one network is equal to the normalized input (or output) admittance of the other and therefore the reflection coefficients have equal magnitudes and opposite signs. The voltage across the load of one

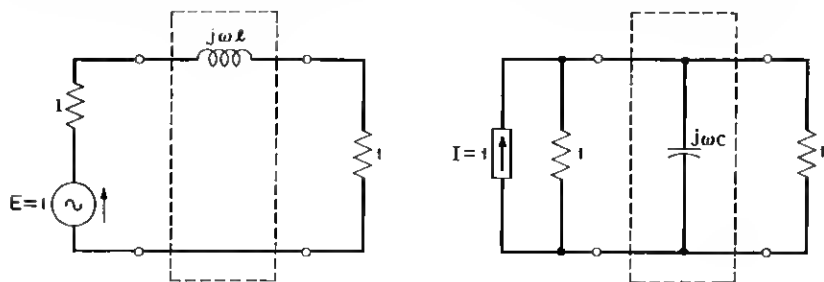


Fig. 12 — An example of dual circuits.

network is equal to the current flowing into the load of the other and therefore the transmissions are the same. Thus, regardless of the frequency, they are oppositely reflective with respect to each other. Since the above explanation is quite general, the theorem is proved.

Now, consider two closely spaced, thin and short parallel conductors and let us connect identical load Z_o 's (real) at each end of the conductors. When the even mode is fed from one end of the conductors, the conductors should represent a lumped L . For the odd mode, the capacitance C between the conductors becomes effective while the currents in the two conductors cancel each other, giving no inductance effect. Therefore, from Theorems I and II this kind of circuit can work as a lumped constant directional coupler. In order to satisfy the required dual property, L/C has to be equal to Z_o^2 . α and β can be calculated from Fig. 12 where the normalized inductance l corresponds to $2L/Z_o$. The coefficient 2 appears here because two loads Z_o are connected in parallel for the even mode to feed current to L .

If one realizes that coupled transmission lines exhibit a dual property to even and odd modes, the theory of the distributed coupler can be developed in a similar fashion. In the limiting case where the coupling factor k approaches unity and the electrical length θ of the coupled region approaches zero, the distributed coupler can be considered as a lumped constant directional coupler.

Although we have not discussed multisection directional couplers, they are useful in obtaining wider bandwidths. The following theorem serves as a guiding principle for constructing such couplers.

Theorem III: The oppositely reflective network of a cascade connection of two-port networks is equivalent to the cascade connection of the oppositely reflective two-port networks (provided that, for the comparison of opposite reflectivity, the same resistance is used for reference at each corresponding reference plane).

This theorem, together with Theorem I, guarantees that when several directional couplers of the type discussed above are connected in cascade, the resulting structure still works as a directional coupler. For the proof, let us first consider a cascade connection of two two-port networks. Using a signal flow graph similar to Fig. 10, the scattering matrix components of the cascade connection are given by

$$S_{11} = S_{11}(1) + \frac{S_{21}(1)S_{12}(1)S_{11}(2)}{1 - S_{22}(1)S_{11}(2)}$$

$$S_{12} = \frac{S_{12}(1)S_{12}(2)}{1 - S_{22}(1)S_{11}(2)}$$

$$S_{21} = \frac{S_{21}(1)S_{21}(2)}{1 - S_{22}(1)S_{11}(2)}$$

$$S_{22} = S_{22}(2) + \frac{S_{22}(1)S_{21}(2)S_{12}(2)}{1 - S_{22}(1)S_{11}(2)}.$$

If we change the signs of $S_{11}(1)$, $S_{22}(1)$, $S_{11}(2)$, and $S_{22}(2)$ then the signs of S_{11} and S_{22} change but S_{12} and S_{21} remain the same. This means that the theorem is true for two networks in cascade. Next, let us increase the number of networks to 3, and first consider No. 1 network as one network and the cascade connection of No. 2 and No. 3 as the other. Then the application of the above proof for two networks shows that the opposite reflective network of the cascade connection of No. 1, No. 2 and No. 3 is equivalent to the cascade connection of the oppositely reflective network of No. 1 and that of No. 2 and No. 3 in cascade. Since another application of the above proof to the cascade connection of No. 2 and No. 3 shows that the oppositely reflective network of No. 2 and No. 3 in cascade is equivalent to the cascade connection of the oppositely reflective networks of No. 2 and No. 3, the theorem is proved for the case of three networks in cascade. Since this procedure of increasing the number of networks can be continued indefinitely, the proof of the theorem is completed.

REFERENCES

1. Engelbrecht, R. S., and Kurokawa, K., A Wideband Low Noise L-Band Balanced Transistor Amplifier, *Proc. IEEE*, 53, No. 3, March, 1965, pp. 237-247.
2. Jones, E. M. T., and Bolljahn, J. T., Coupled Strip Transmission Line Filters and Directional Couplers, *IRE Trans. MTT.*, 4, April, 1956, pp. 75-81.
3. Kurokawa, K., Actual Noise Measure of Linear Amplifiers, *Proc. IRE*, 49, Sept., 1961, pp. 1391-1397.
4. Experiments conducted by A. L. Stillwell and F. J. D'Alessio.